Molecular Phylogeny of Genus *Horaglanis* (Indian Blind Catfishes) Within the Family Clariidae

Annam Pavan-Kumar  
Fish Genetics and Biotechnology Division, ICAR-Central Institute of Fisheries Education, Yari Road, Panch Marg, Versova, Mumbai – 400061  
Email: pavankumar@cife.edu.in
Abstract

Background: India is one of the mega biodiverse countries with a large number of endemic freshwater fishes. Recently, species of genus *Horaglanis* (family: Clariidae) have been reported from the southern part of India. Due to their unique morphological adaptations, these enigmatic species have been subjected to phylogenetic studies to understand the evolution of adaptive traits. Further, the taxonomic status of these species has not been verified using molecular data. Methods: In the present study, secondary data i.e. reported sequences of mitochondrial cytochrome *c* oxidase subunit I gene was used to estimate the genetic divergence values for 14 species of the family Clariidae. Phylogenetic trees were reconstructed using maximum parsimony, maximum likelihood and Bayesian Inference methods. Results: The average genetic divergence value among genera *Clarias-Clariallabes-Platyallabes-Dolichallabes-Gymnallabes-Tanganikallabes-Channallabes* was 0.082 ± 0.01. However, these genera showed an average divergence value of 0.296±0.02 with genus *Horaglanis*. In all tree topologies, species of the genus *Horaglanis* formed a basal group to all other clariids. Discussion: The higher genetic divergence value between genus *Horaglanis* and other genera of Clariidae family suggested that genus *Horaglanis* may belong to a separate sub family. Based on phylogenetic trees, it is evident that species of *Horaglanis* might have originated early in the evolution of Clariids than other species.

Keywords: Troglomorphic fish, Molecular phylogeny, Cytochrome *c* oxidase subunit I, Clariidae, Horaglanis
Introduction

India is one of the mega biodiverse countries with a large number of endemic freshwater fishes. A total of 946 freshwater fish species (endemic species: 195 nos) have been reported from India (FishBase, 2017). Among different ecosystems, caves and other subterranean habitats are one of the most challenging environments for fishes. Nevertheless, bony fishes are the only vertebrate group that has been successfully colonized these habitats (Soares & Niemiller, 2013). Due to their unique morphological adaptations such as lack of eyes and pigments, these cave fishes / hypogean fishes have been used as a model organism to study the evolution of adaptive traits (Romeo & Paulson, 2001).


Among these, species of Horaglanis (order: Siluriformes) are endemic to India and commonly known as Indian blind catfish. Earlier, Menon (1951) described blind catfishes from India and placed them in family Clariidae under genus Horaglanis. Later, Jayaram (2006; 2010) categorized genus Horaglanis under a new subfamily Horaglanidinae. However, the evolutionary relationship/phylogeny of genus Horaglanis has not been investigated using either morphological or molecular data. Further, the Phylogenetic analysis will provide new insights into the taxonomical position and character evolution of this species. Among the molecular markers, phylogenetically informative gene mitochondrial cytochrome c oxidase subunit I (COI “barcode gene”) has been successfully used for metazoan identification and species delimitation (Hebert et al., 2004, Hajibabaei, 2012). It resulted in a global initiative to develop a comprehensive DNA barcode database (Barcode of Life Database: BOLD) including species specific markers (COI gene) along with their taxonomic details. At present, this database contains taxonomic information for 11256 fish species (Ratnasingam and Hebert, 2007). The COI sequences of this database could be used to infer phylogenetic relationship between species at lower taxonomic level (i.e. within the family). Previously several researchers have used the DNA sequences available from databases and resolved the phylogenetic relationship between different taxon through in silico approach (Dangre et al., 2009; Vélez-Zuazo and Agnarsson, 2011).
With this background, the present study was aimed to investigate the phylogenetic position of genus *Horaglanis* within Clariidae family using mitochondrial COI gene sequences available from public databases.

**Materials and Methods**

**Sequence data mining and analysis**

The public data portal of BOLD and core nucleotide database of GenBank were searched for COI sequences of Clariidae fishes. A total of 39 COI sequences representing 8 genera, 14 species were downloaded from databases (Table 1). In BOLD / GenBank database, the species name of genus *Horaglanis* was not defined and reported as *Horaglanis* species. *Heteropneustes fossilis* (Heteropneustidae family) was taken as an outgroup for phylogenetic studies. COI sequences with a minimum length of 600 bp, with no missing nucleotides or gaps, were included for analyses. Open reading frame was predicted for all the COI sequences using NCBI ORF finder to confirm the lack of NUMTs (nuclear DNA originating from mitochondrial DNA sequences, Zhang & Hewitt, 1996).

**Model of evolution and phylogeny reconstruction methods**

The sequences were aligned using ClustalW and genetic divergence values were calculated using Kimura two parameter (K2P) distance model implemented in MEGA V.5.2 (Tamura et al., 2011). The size of the alignment, number of variable sites, number of parsimony informative sites among all taxa were calculated using MEGA V.5.2 software. Chi-square ($\chi^2$) test was used to test the base compositional evenness and stationarity among taxa, implemented in PAUP* v.4.0b10 (Swofford, 2003) software. J MODELTEST v.0.1.1 (Posada, 2008) was used to estimate the most likely model of sequence evolution for Maximum likelihood (ML) and Bayesian Inference (BI) methods. Based on likelihood values and the Akaike information Criterion (AIC), most likely sequence evolutionary models were selected.

Distance based (K2P models), Maximum Parsimony and Maximum likelihood analyses were performed by PAUP* v.4.0b10 (Swofford, 2003). For Maximum Parsimony analysis, all characters were treated as unordered and un-weighted. A heuristic search for the most parsimonious trees was conducted with random addition sequence (100 replicates) and tree bisection–reconnection (TBR) branch-swapping. A majority-rule consensus tree was constructed with 1000 bootstrap resampling (Felsenstein, 1985) to assess support of relationships. MR BAYES v3.1.2 (Huelsenbeck & Ronquist, 2003) was used for Bayesian inference. For all
analyses, the prior probability of a flat Dirichlet distribution for the substitution rates and
stationary nucleotide frequencies was used under the assumption of no prior knowledge. The
Bayesian analysis was run using the Metropolis coupled Markov Chain Monte Carlo (MCMC)
algorithm from randomly generated starting trees for 5 million generations with sampling every
1000 generations. Two runs were performed simultaneously in which there were three heated
chains and one cold chain, each with a temperature parameter of 0.2. The first 25% of sampled
trees were discarded as burn-in.

Results and Discussion
Mitochondrial COI genes with 602 nucleotide characters were aligned to the
homologous position. The number of conserved, variable and parsimony informative
nucleotides is 374, 228 and 218, respectively. These values show that the COI gene has sufficient
phylogenetic signal to infer the evolutionary relationship. The average nucleotide frequencies for
COI gene among the species were A = 26.9, T=29, G=17.1 and C= 26.9% with a GC content
of 44%. The GC content of COI gene region is slightly lesser than previous reports for teleosts
(Ward et al., 2005). The average GC content values at codon 1\textsuperscript{st}, 2\textsuperscript{nd}, and 3\textsuperscript{rd} base positions are
55.4, 42.5 and 34.2\%, respectively. The average transitional pairs ($S_{t}= 50$) were more than the
transversional pairs ($S_{v}=29$) with an average ratio of 1.75. It showed that the sequences were not
saturated and useful for phylogenetic analysis. Chi-square ($\chi^2$) test applied for detecting
homogeneity of base composition across the taxa indicated that there was no significant variation
($P = 0.99$) in AT/GC content among species.

The species name for the COI sequences of Horaglanis genus were not described in
NCBI GenBank /BOLD databases. To confirm whether these COI sequences belongs to the
same species or not, genetic divergence values were estimated and the results revealed a genetic
distance value of “0” suggesting that the four specimens indeed belong to the same species of
genus Horaglanis.

The average genetic divergence value among genera Clarias, Clariallabes, Platyallabes,
Dolichallabes, Gymnallabes, Tanganikallabes and Channallabes was 0.082 ± 0.01. However, the average
genetic divergence value between these genera and genus Horaglanis was 0.296 ± 0.12 (Table 2).
In the present study, degree of increase in genetic distance value (3X) among genera is more than
the previous reports (Ward et al., 2005; Lakra et al., 2011). Based on the high genetic divergence
value between genus Horaglanis and other genera of family Clariidae, genus Horaglanis could be
grouped under separate subfamily “Horaglanidinae”. Nucleotide diagnostic characters of COI
sequences exclusive to genus *Horaglanis* were identified by comparing with other Clariidae species. Simple nucleotide diagnostic characters were identified in 56 informative sites. Among these, 26 characters correspond to transitions and remaining 30 sites are transversional changes (Supplementary Figure 1). These diagnostic characters could help in species of the genus *Horaglanis* identification. Generally, nucleotide diagnostic characters are useful for discriminating closely related species (Wong et al., 2009). However, in this study we used the only one reported *Horaglanis* species, but we believe that in view of discovering more number of different species of *Horaglanis* (Babu, 2012), this information will help in identifying/discriminating the closely related *Horaglanis* species.

The phylogenetic relationship among Clariidae species was estimated by Neighbor-joining, Maximum Parsimony, Maximum likelihood (TIM1+I+G model) and Bayesian Inference (GTR+I+G model) methods. Maximum parsimony method yielded most parsimonious tree with a length of 583. The consistency, homoplasy and retention indexes of the tree are 0.58, 0.41 and 0.79, respectively.

All methods yielded similar topologies with a slight difference in Bootstrap values (Supplementary figure 2-4). Some of the tree topologies displayed soft polytomies at internal nodes and this could be due to the usage of a single gene. However, in all tree topologies, species of the genus *Horaglanis* formed the basal group to all other Clariids (distinct lineages) and the branch length was relatively high (Fig. 1). It showed that the species might have originated early in the evolution of Clarids than other species. Previous studies have compared the cranial osteology of species of *Horaglanis* with other genera of Clariidae fishes and reported Genus *Horaglanis* as a primitive Silurid fishes (Bhimachar, 1933, Mercy & Pillai, 2001, Mercy et al., 2001). Bhimachar (1933) reported that species of *Horaglanis* would be a connecting link between family Clarids and Bagrids. Based on the COI divergence and phylogeny analyses, the present study also supports that the genus *Horaglanis* could be grouped into a separate subfamily. However, it needs further studies with more number of molecular markers and species to understand the evolutionary relationship of these enigmatic species. The phylogenetic tree also revealed the paraphyletic nature of species of *Clarias* as reported by previous workers (Devaere et al., 2005, 2007; Jansen et al., 2006; Mwita & Nkwengulila, 2008).

In conclusion, this study provides preliminary information on the phylogenetic position of the genus *Horaglanis* within the clariidae family based on the available molecular marker
information. Further, this study identified nucleotide diagnostic characters exclusive to the genus *Horaglanis*. 
References


Table 1: List of species along with GenBank / BOLD BIN numbers

<table>
<thead>
<tr>
<th>S.No</th>
<th>Species (number of specimen)</th>
<th>Characteristics</th>
<th>GenBank accession / Bold ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td><em>Channallabes apus</em> (Günther, 1873) (4)</td>
<td>Endemic to Africa, anguilliform shape, reduced suprabranchial organ (Fishbase)</td>
<td>AMNH001-09, AMNH008-09, AMNH0159-09, AMNH01321-10</td>
</tr>
<tr>
<td>2.</td>
<td><em>Gymnallabes nops</em> Roberts &amp; Stewart, 1976 (1)</td>
<td>Endemic to Africa, anguilliform shape, reduced suprabranchial organ (fishbase)</td>
<td>AMNH0121-09</td>
</tr>
<tr>
<td>3.</td>
<td><em>Clarias teugelsi</em> Ferraris, 2007 (3)</td>
<td>Endemic to Africa, elongate body shape, suprabranchial organ present (Fishbase)</td>
<td>AMNH0125-09, AMNH0126-09, AMNH01498-12</td>
</tr>
<tr>
<td>4.</td>
<td><em>Dolichallabes microphthalmus</em> Poll, 1942 (3)</td>
<td>Endemic to Africa, anguilliform shape, reduced suprabranchial organ (Fishbase)</td>
<td>AMNH0121-09</td>
</tr>
<tr>
<td>5.</td>
<td><em>Platyallabes tihoni</em> (Poll, 1944) (2)</td>
<td>Endemic to Africa, intermediate shape between anguilliform and fusiform, reduced suprabranchial organ (Devaere et al., 2005)</td>
<td>AMNH0125-09, AMNH0126-09, AMNH01498-12</td>
</tr>
<tr>
<td>6.</td>
<td><em>Tanganikallabes mortiauxi</em> Poll, 1943 (2)</td>
<td>Endemic to deeper waters in Lake Tanganyika (East Africa), Fusiform shape, reduced suprabranchial organ (Write and Bailey 2012)</td>
<td>AMNH0129-09, AMNH01210-09</td>
</tr>
<tr>
<td>7.</td>
<td><em>Clarias gabonensis</em> Günther, 1867 (5)</td>
<td>Endemic to Africa, anguilliform shape, developed suprabranchial organ (Fishbase)</td>
<td>AMNH0122-10, AMNH0124-10, AMNH01494-12, AMNH01495-12, BAFEN135-10</td>
</tr>
<tr>
<td>8.</td>
<td><em>Clarias angolensis</em> Steindachner, 1866 (1)</td>
<td>Africa; Lower and middle Congo river, fusiform, developed suprabranchial organ</td>
<td>HM880232</td>
</tr>
<tr>
<td>9.</td>
<td><em>Clarias teijsmanni</em> Bleeker, 1857 (1)</td>
<td>Endemic to Asia, anguilliform shape (Fishbase)</td>
<td>JN646093</td>
</tr>
<tr>
<td>10.</td>
<td><em>Clarias gariepinus</em> (Burchell, 1822) (5)</td>
<td>Native of Africa, fusiform Suprabranchial organ (Fishbase)</td>
<td>JQ699199-203</td>
</tr>
<tr>
<td>11.</td>
<td><em>Clarias dussumieri</em> Valenciennes, 1840 (5)</td>
<td>Native to Asia, Fusiform, slightly reduced suprabranchial organ (Fishbase)</td>
<td>JQ699209-213</td>
</tr>
<tr>
<td>12.</td>
<td><em>Clarias camerunensis</em> Lönnberg, 1895 (1)</td>
<td>Endemic to Africa, anguilliform, developed suprabranchial organ (fishbase)</td>
<td>HM882808</td>
</tr>
<tr>
<td>13.</td>
<td><em>Clarias jaensis</em> Boulenger, 1909 (2)</td>
<td>Africa, anguilliform, developed suprabranchial organ (Fishbase)</td>
<td>BAFEN138-10, BAFEN139-10</td>
</tr>
<tr>
<td>14.</td>
<td><em>Horaglanis</em> species (4)</td>
<td>Asia, anguilliform shape, reduced suprabranchial organ (Fishbase)</td>
<td>HE819391-94</td>
</tr>
<tr>
<td>15.</td>
<td><em>Heteropneustes fossilis</em> (4)</td>
<td></td>
<td>GQ466396-97</td>
</tr>
</tbody>
</table>
Table 2. COI Average genetic divergence values (K2P method) among Clariidae species

<table>
<thead>
<tr>
<th>Species</th>
<th>C.apus</th>
<th>G.nops</th>
<th>C.teugelsi</th>
<th>D.microphthalmus</th>
<th>P.tiboni</th>
<th>T.mortiauxi</th>
<th>C.gabonensis</th>
<th>C.teijsmanni</th>
<th>C.gariepinus</th>
<th>C.dussumieri</th>
<th>C.camerunensis</th>
<th>C_jaensis</th>
<th>Horaglanis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channallabes apus</td>
<td>0.013</td>
<td>0.013</td>
<td>0.013</td>
<td>0.012</td>
<td>0.008</td>
<td>0.011</td>
<td>0.015</td>
<td>0.013</td>
<td>0.013</td>
<td>0.013</td>
<td>0.014</td>
<td>0.011</td>
<td>0.026</td>
</tr>
<tr>
<td>Gymnallabes nops</td>
<td>0.095</td>
<td>0.014</td>
<td>0.011</td>
<td>0.014</td>
<td>0.014</td>
<td>0.014</td>
<td>0.016</td>
<td>0.016</td>
<td>0.015</td>
<td>0.014</td>
<td>0.013</td>
<td>0.013</td>
<td>0.027</td>
</tr>
<tr>
<td>Clariallabes teugelsi</td>
<td>0.102</td>
<td>0.108</td>
<td>0.014</td>
<td>0.012</td>
<td>0.014</td>
<td>0.014</td>
<td>0.017</td>
<td>0.016</td>
<td>0.016</td>
<td>0.015</td>
<td>0.015</td>
<td>0.014</td>
<td>0.026</td>
</tr>
<tr>
<td>Dolichallabes microphthalmus</td>
<td>0.089</td>
<td>0.059</td>
<td>0.105</td>
<td>0.014</td>
<td>0.013</td>
<td>0.013</td>
<td>0.016</td>
<td>0.015</td>
<td>0.015</td>
<td>0.015</td>
<td>0.014</td>
<td>0.012</td>
<td>0.026</td>
</tr>
<tr>
<td>Platyallabes tiboni</td>
<td>0.077</td>
<td>0.110</td>
<td>0.127</td>
<td>0.113</td>
<td>0.012</td>
<td>0.012</td>
<td>0.016</td>
<td>0.015</td>
<td>0.013</td>
<td>0.015</td>
<td>0.015</td>
<td>0.012</td>
<td>0.026</td>
</tr>
<tr>
<td>Tanganikallabes mortianzi</td>
<td>0.038</td>
<td>0.103</td>
<td>0.086</td>
<td>0.091</td>
<td>0.077</td>
<td>0.010</td>
<td>0.015</td>
<td>0.014</td>
<td>0.013</td>
<td>0.014</td>
<td>0.011</td>
<td>0.026</td>
<td></td>
</tr>
<tr>
<td>Clarias gabonensis</td>
<td>0.064</td>
<td>0.101</td>
<td>0.111</td>
<td>0.091</td>
<td>0.079</td>
<td>0.058</td>
<td>0.015</td>
<td>0.015</td>
<td>0.013</td>
<td>0.014</td>
<td>0.013</td>
<td>0.027</td>
<td></td>
</tr>
<tr>
<td>Clarias teijsmanni</td>
<td>0.132</td>
<td>0.130</td>
<td>0.153</td>
<td>0.136</td>
<td>0.132</td>
<td>0.121</td>
<td>0.016</td>
<td>0.015</td>
<td>0.013</td>
<td>0.014</td>
<td>0.016</td>
<td>0.024</td>
<td></td>
</tr>
<tr>
<td>Clarias gariepinus</td>
<td>0.098</td>
<td>0.133</td>
<td>0.131</td>
<td>0.147</td>
<td>0.118</td>
<td>0.117</td>
<td>0.132</td>
<td>0.015</td>
<td>0.013</td>
<td>0.015</td>
<td>0.015</td>
<td>0.028</td>
<td></td>
</tr>
<tr>
<td>Clarias dussumieri</td>
<td>0.096</td>
<td>0.126</td>
<td>0.141</td>
<td>0.116</td>
<td>0.104</td>
<td>0.104</td>
<td>0.125</td>
<td>0.123</td>
<td>0.015</td>
<td>0.014</td>
<td>0.027</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clarias camerunensis</td>
<td>0.106</td>
<td>0.112</td>
<td>0.112</td>
<td>0.118</td>
<td>0.120</td>
<td>0.104</td>
<td>0.134</td>
<td>0.095</td>
<td>0.113</td>
<td>0.013</td>
<td>0.028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clarias jaensis</td>
<td>0.068</td>
<td>0.095</td>
<td>0.112</td>
<td>0.099</td>
<td>0.085</td>
<td>0.071</td>
<td>0.130</td>
<td>0.119</td>
<td>0.108</td>
<td>0.098</td>
<td>0.026</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horaglanis species</td>
<td>0.298</td>
<td>0.301</td>
<td>0.317</td>
<td>0.308</td>
<td>0.303</td>
<td>0.301</td>
<td>0.314</td>
<td>0.280</td>
<td>0.316</td>
<td>0.308</td>
<td>0.321</td>
<td>0.306</td>
<td></td>
</tr>
</tbody>
</table>

*Below diagonal distance values; above diagonal standard error values for corresponding distance value.
Fig.1. Bayesian inference based phylogeny depicting relationship among clariidae species
Sup. Fig. 1. Nucleotide diagnostic characters specific to species of *Horaglanis*
Sup. Fig. 2. Neighbor-Joining tree of Clariidae species constructed by COI gene
Sup. Fig.3. Clariidae species phylogeny tree constructed by COI gene using Maximum Parsimony method.
Sup. Fig.4. Clariidae Species phylogeny tree constructed by Maximum Likelihood method.